AMD Presentation For Linux Kernel Summit

Richard A. Brunner
AMD Fellow

July 2005

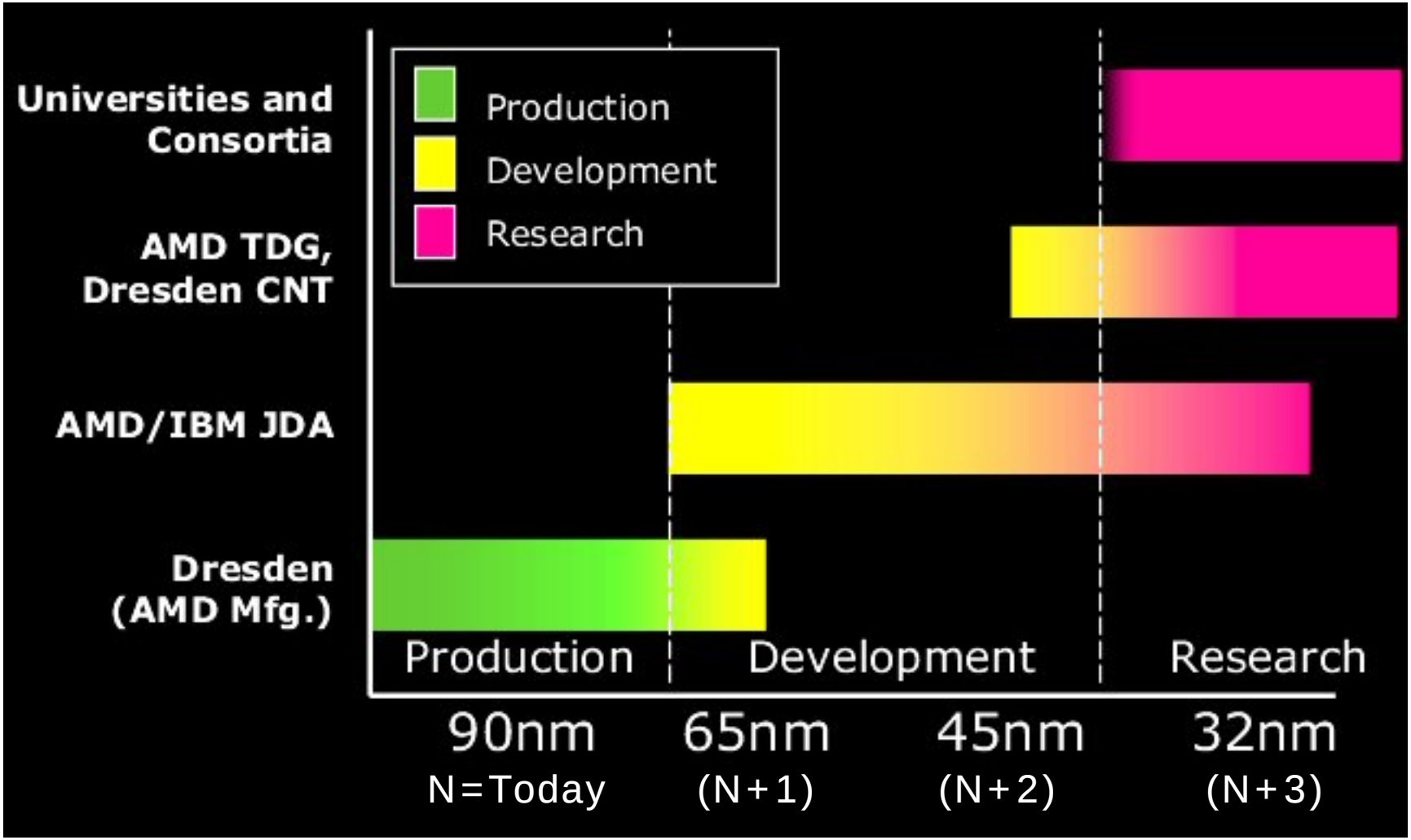
Progress Report

Linux is moving into the mainstream with AMD64 technology

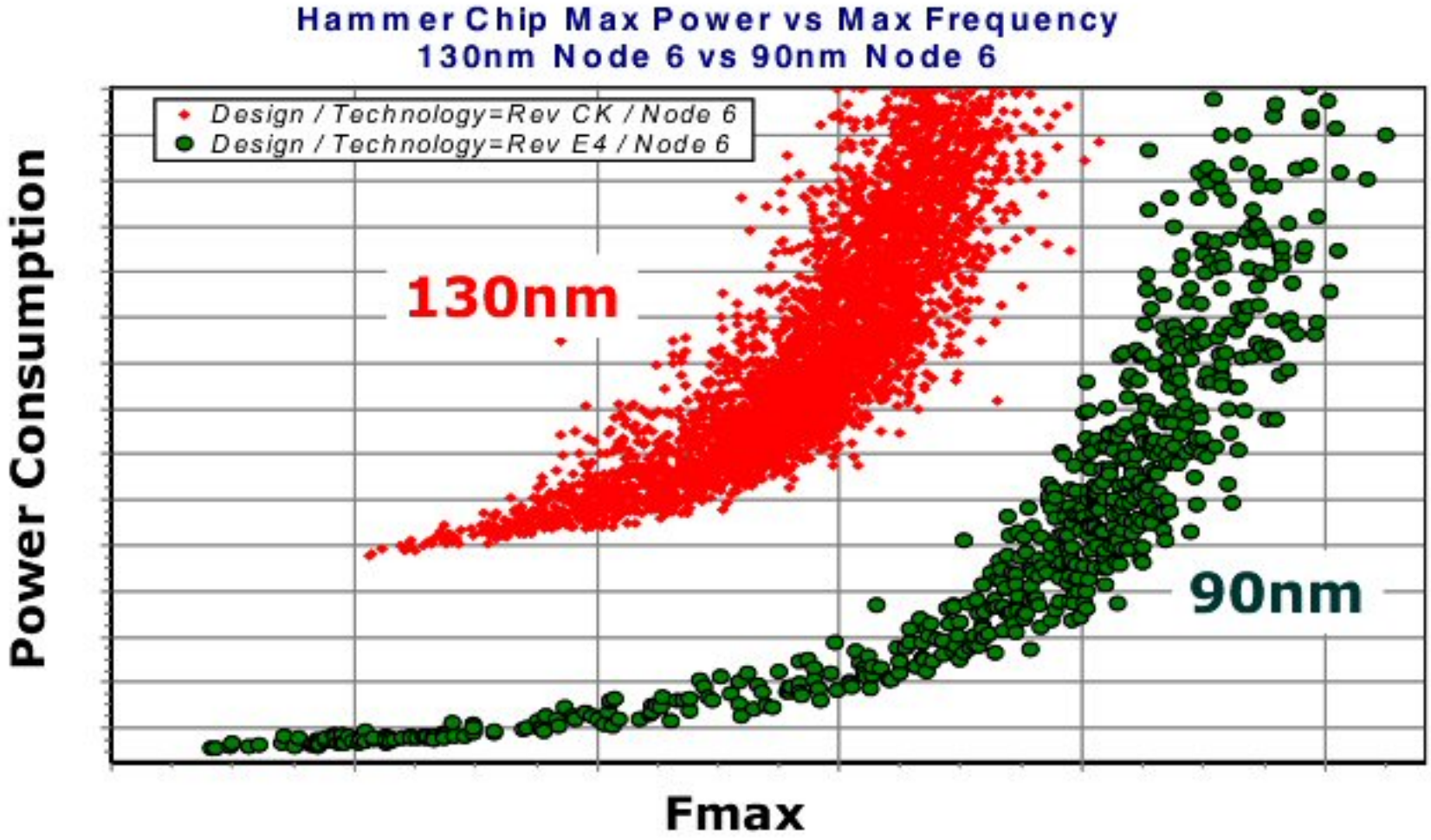
- AMD64+Linux is penetrating deeper into the Data Center
  - Demanding mainframe/UNIX functionality: 64-bit, NUMA, multi-core, and virtualization
  - Requiring solutions to infrastructure issues: more power management, security, and manageability
  - Requesting innovation without disruption: evolution as opposed to revolution - need to maintain compatibility and stability
  - Using servers and workstations as the proving ground: Linux must do well in these area before they move to Linux on the desktop
- AMD continues the trend of openly providing early technical information on our products to the developer community for feedback.

AMD's Technology Roadmap

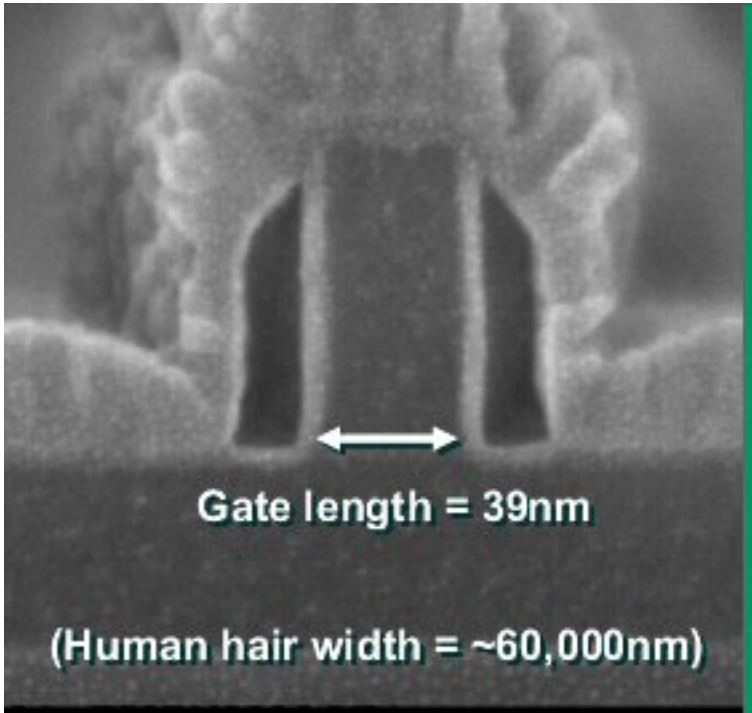
Technology Roadmap



130nm to 90nm Performance & Power



65nm Progress



High-performance 65nm technology SOI transistor with strain engineering

Achieved our 65nm SRAM milestones

SRAM test vehicle yields are exceeding our development plans

Transistor and interconnect development on schedule

Yield metrics exceeding development plan
3rd generation of strained-silicon
Addition of Nickel Silicide
4th generation of low-k dielectric stack

On schedule for development process installation in Fab 36 in mid-2005

AMD is optimizing Fab 36 for smooth transition of 65nm technology

Post-45nm Research Begun and Processing

Key candidates:

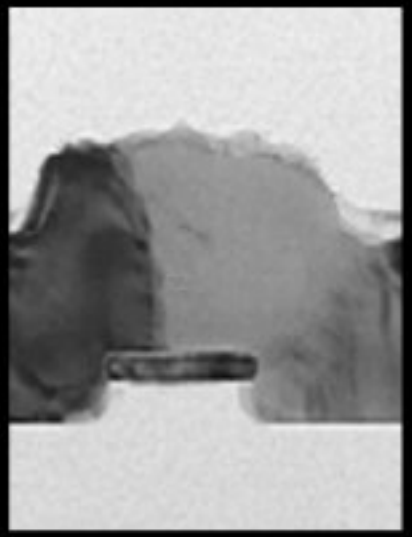
- Ultra-thin, fully-depleted (FDSOI)
- Metal gate (NiSi)
- Intrinsic channels
- Multi-gate 3D structure
- Strain engineering

Expected advantages:

- Lower leakage (Off current and gate leakage)
- Higher drive currents
- Faster switching
- Exceeded ITRS performance projected for 2009

AMD Metal-Gate FDSOI

First Published at SSDM 2003

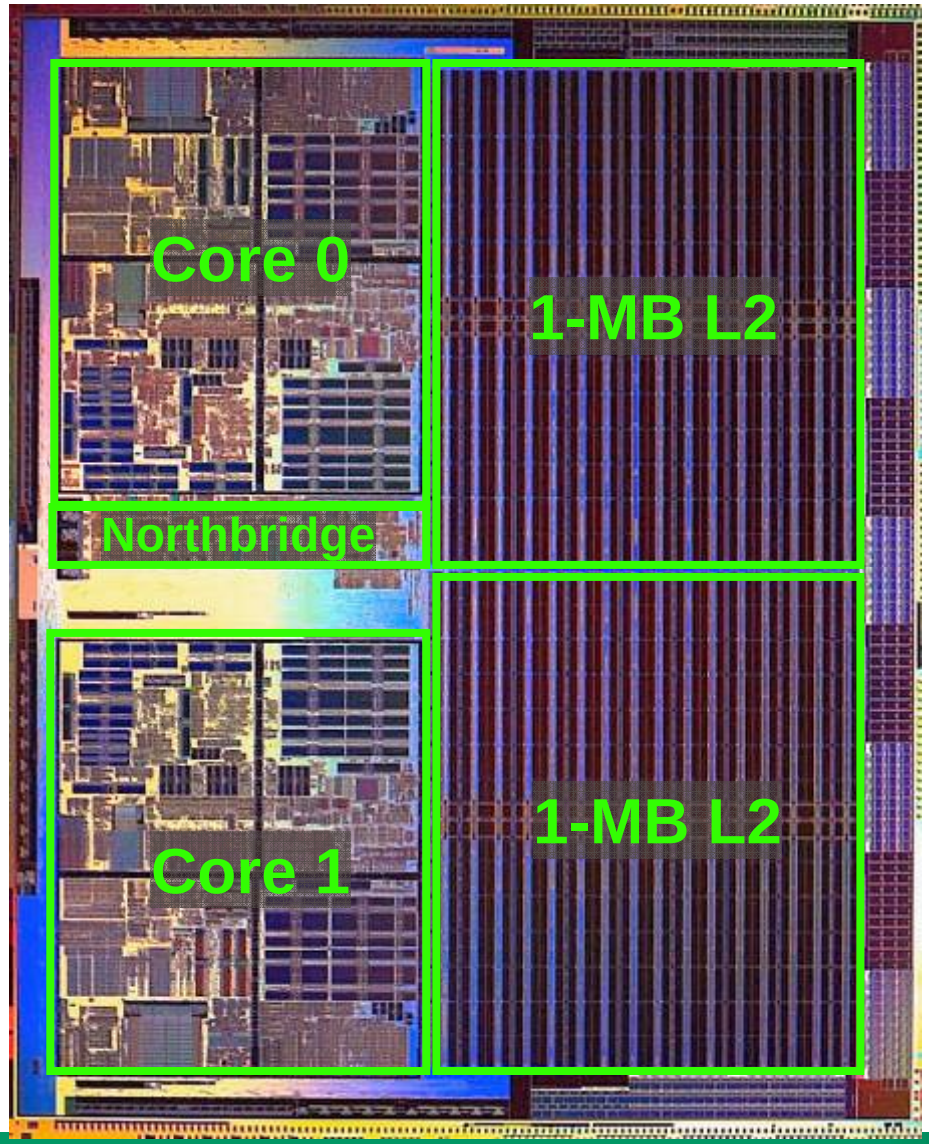


Multi-gate structure Lg=20 nm

Research transistors surpass ITRS 2009 performance projections

AMD's Processor Roadmap

Introducing AMD64 Dual-Core Processor

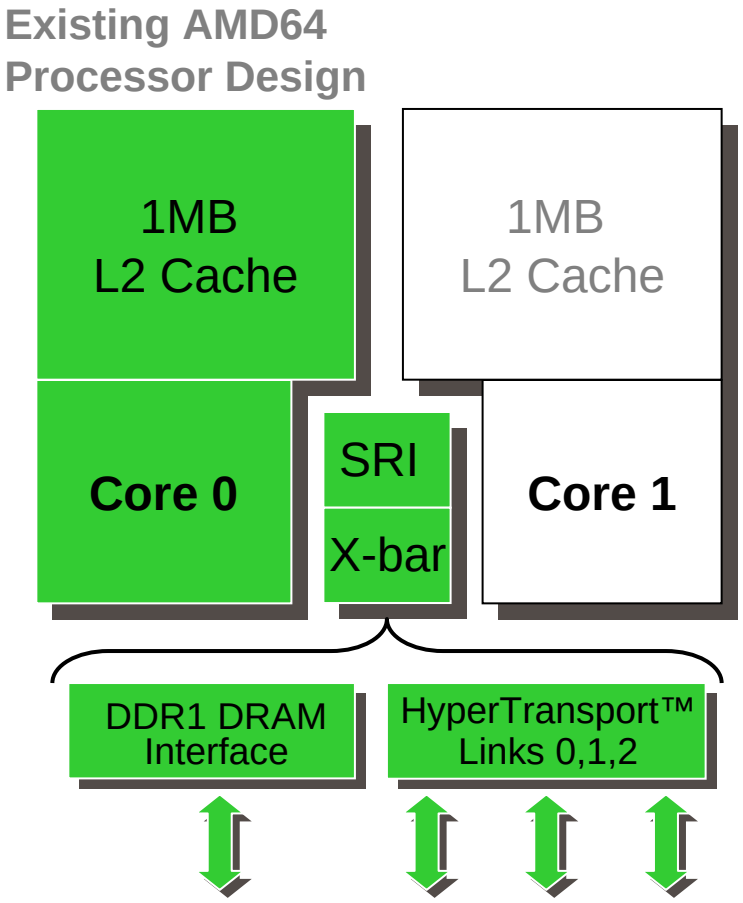


- Two AMD Opteron™ CPU cores on a single die, each with 1MB L2 cache
- 90nm, ~205 million transistors*
  - Approximately same die size as 130nm single-core AMD Opteron processor*
- 95 watt power envelope fits into 90nm power infrastructure
- Retains compatibility with existing 32-bit and 64-bit x86-base software
- Introduced with “K8” Revision E core in April 2005

**Based on current revisions of the design*

Designed From The Start To Add Second Core

- Shared Northbridge
  - 3 HyperTransport™ technology links
  - Dual-channel (128 bit) DDR i/f
- AMD Opteron™ CPU with Direct Connect Architecture was designed as CMP from the start
  - Second port on SRI, request management, two APICs
- Two complete CPU cores
  - SMP model
  - Simpler, less-restrictive programming model than “logical core” approach
  - No need to “pause” one core to give other exclusive use of shared resources



AMD Dual-Core Technology

Desktop

AMD Athlon™ 64 X2 Dual-Core Processor *(Announced June 2005)*		
Model #	Freq	L2 Cache
4800+	2.4 Ghz	1 MB + 1 MB
4600+	2.4 Ghz	512KB + 512KB
4400+	2.2 Ghz	1 MB + 1 MB
4200+	2.2 Ghz	512KB + 512KB

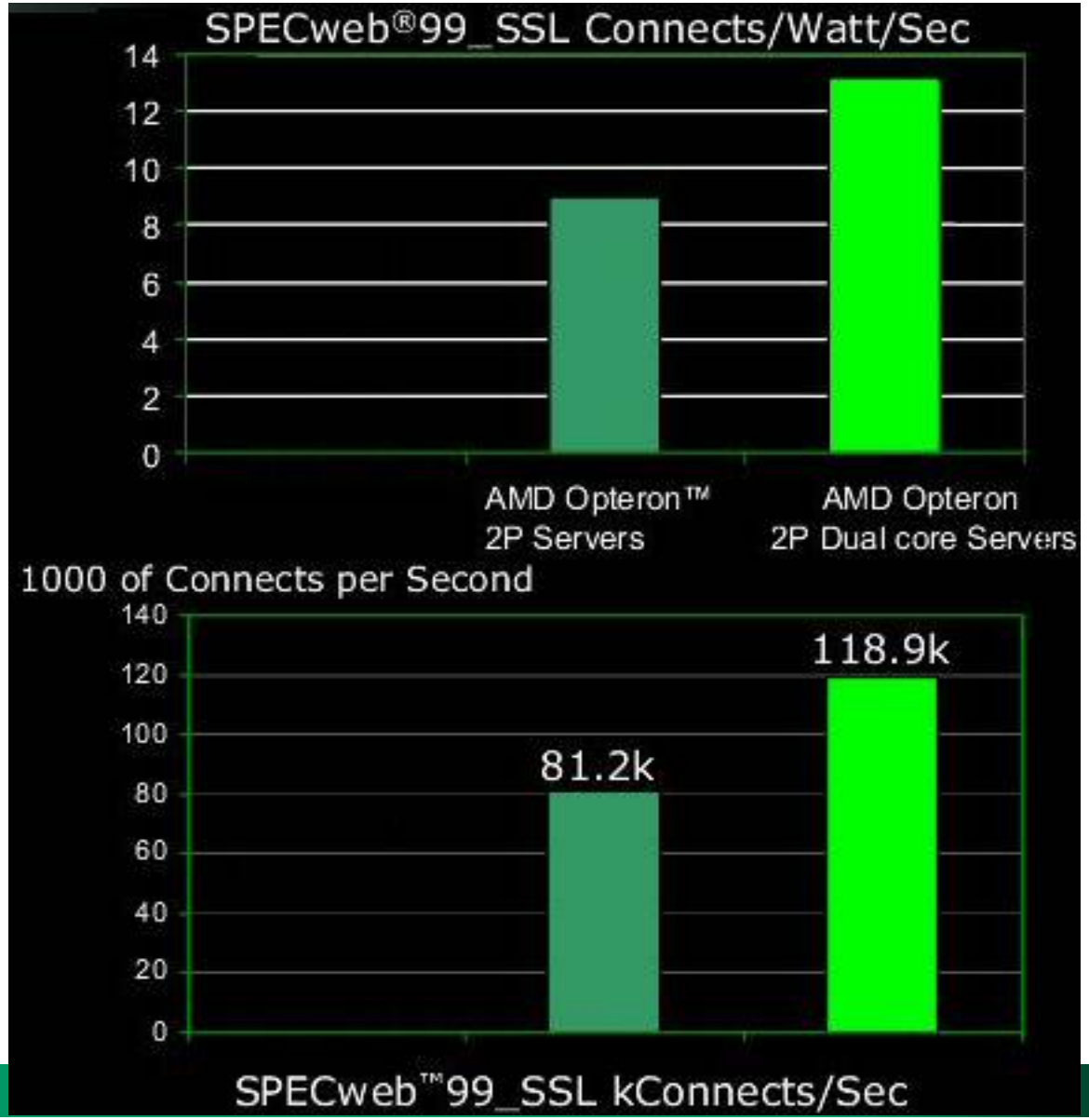
http://www.amd.com/us-en/Processors/ProductInformation/0,,30_118_8796_9240,00.html

Server/Workstation

AMD Opteron™ Processor Dual-Core Models *(Announced on April 21, 2005)*			
Freq	1-way	Up to 2-way	Up to 8-way
1.8 GHz	Model 165	Model 265	Model 865
2.0 GHz	Model 170	Model 270	Model 870
2.2 GHz	Model 175	Model 275	Model 875

http://www.amd.com/us-en/Processors/ProductInformation/0,,30_118_9485_13041%5E13076,00.html

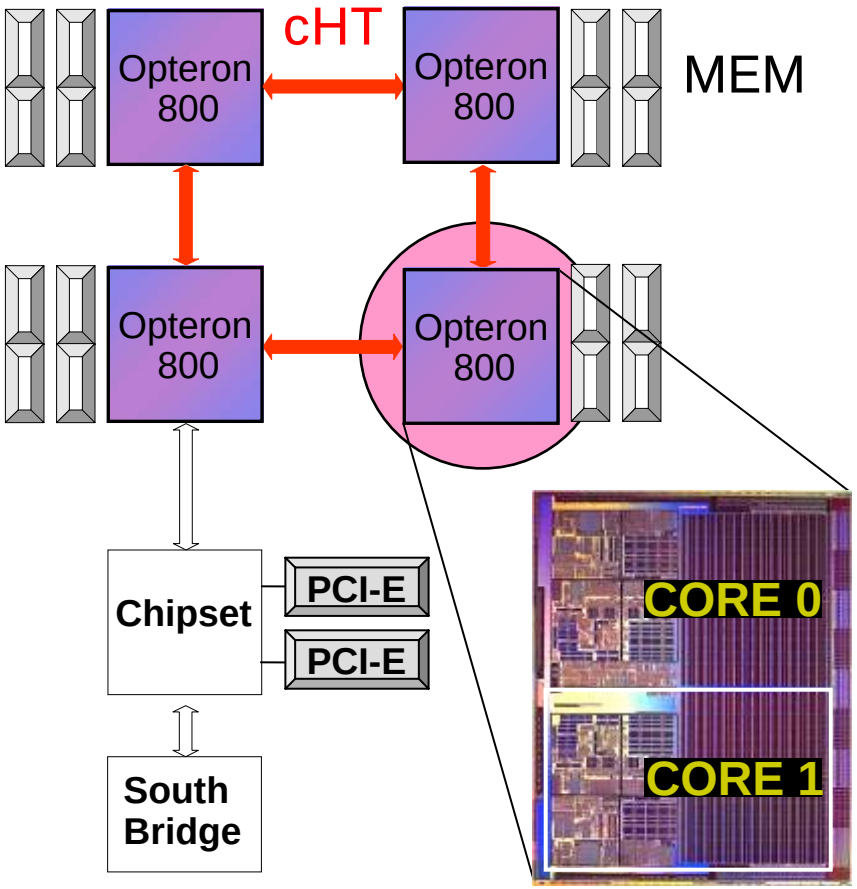
Dual-Core Performance / Watt



- SPECweb® 99_SSL Secure Web Connections Example.
- Data Center rack space and power budgets are often fixed.
- Perf/Watt focus maximizes use of resources.
- Typical 48U Rack has 9KVA of Power.

AMD Direct Connect Architecture + Dual-Core

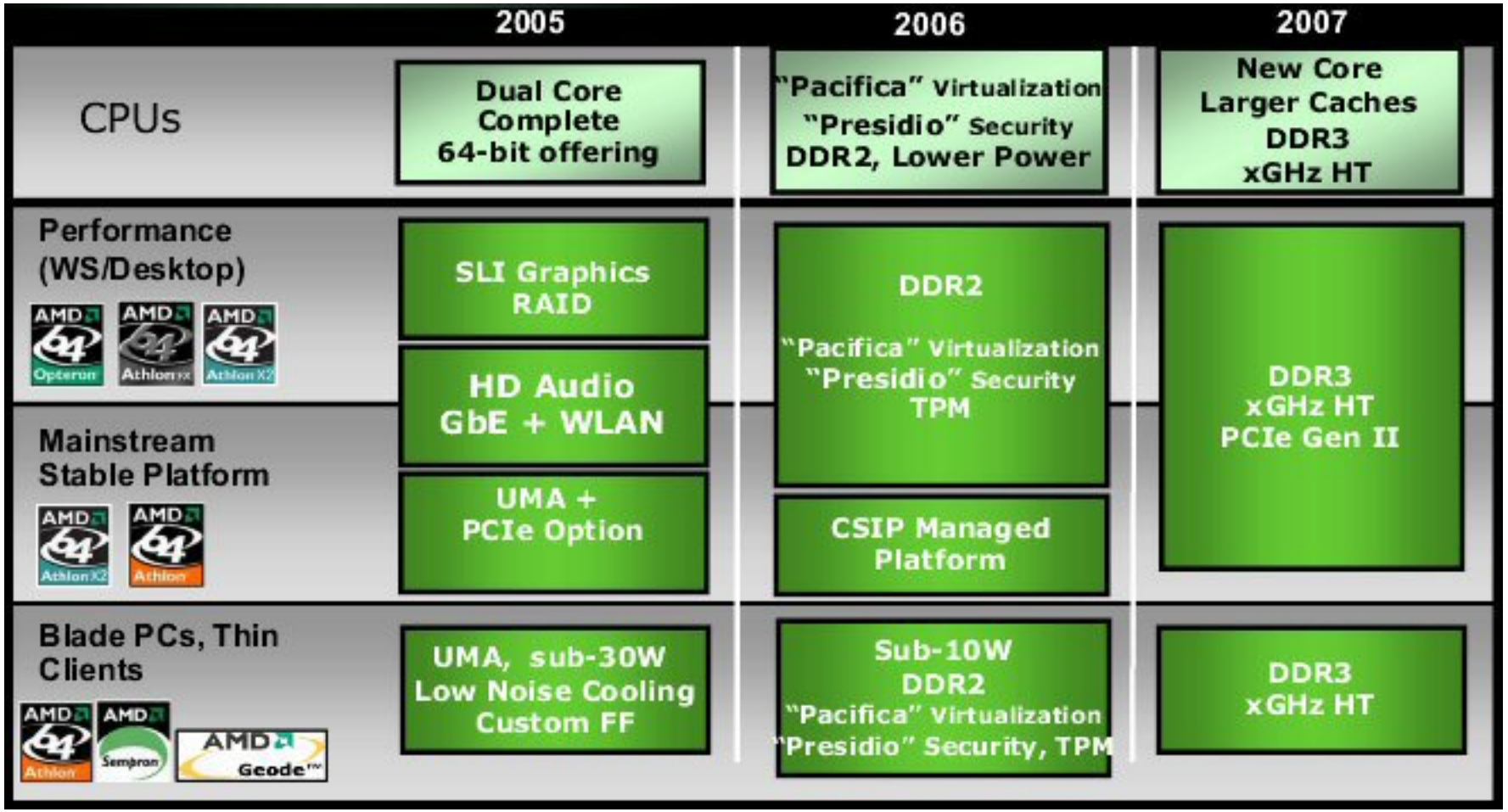
	cHT		MEM	
K8 REV	cHT (MHz)	BW 16x16 1-w/2-w (GB/s)	DDR (MHz)	BW 1-ch/2-ch (GB/s)
CG	HT1-800	3.2/6.4	DDR1-400	3.2/6.4
E	HT1-1000	4.0/8.0	DDR1-400	3.2/6.4
F	HT1-1000	4.0/8.0	DDR2	



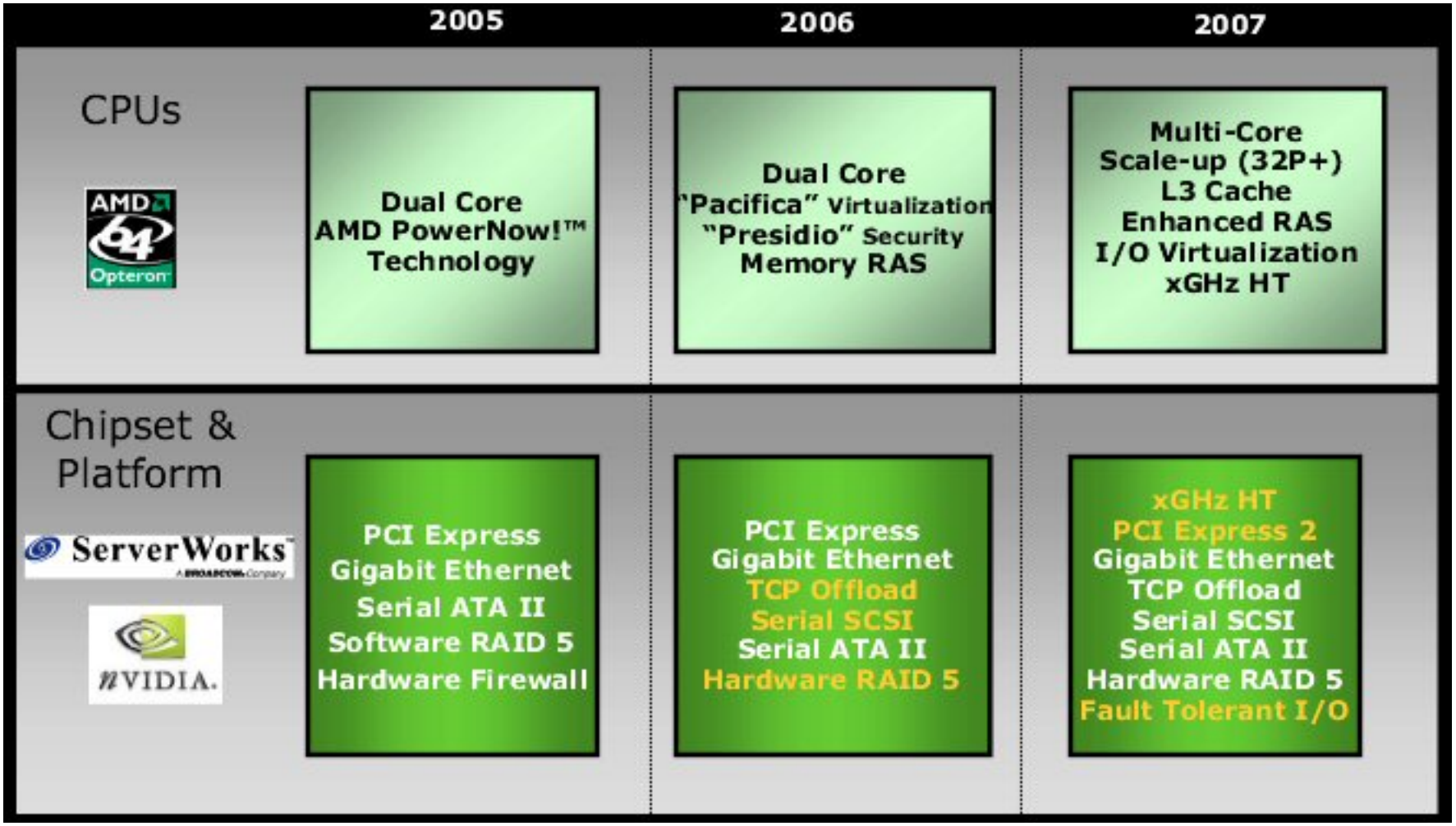
SSE3 Support

- AMD K8 Revision “E” and newer are designed to support SSE3
- Supports SSE3 instructions reported by CPUID.SSE3 feature flag
- Ten new SSE instructions and one new x87 instruction (13 total opcodes).
- Monitor/Mwait planned in 2007
- CMPXCHG16B planned in 2006
- ADDSUB[PD,PS] xmm1, xmm2/m128
  - Provides interleaved packed add and subtract
- FISTTP m16int/m32int/m64int
  - Like FISTP but with forced truncation
- HADD[PD,PS] xmm1, xmm2/m128
  - Horizontal Adds
- HSUB[PD,PS] xmm1, xmm2/m128
  - Horizontal Subtracts
- LDDQU xmm, m128
  - Special 128-bit Unaligned load
- MOV[D,HD,LD]DUP xmm1, xmm2/m64
  - Move and Duplicate some elements

Desktop / Workstation Roadmap



Server / Workstation Roadmap



Planned 2006 Processor Features

- Multi-Core capable
- DDR2 support
- RDTSCP – see next slide
- CMPXCHG16B – compare 16bytes, exchange 16-bytes
- Correctable Machine-Check Exception Thresholding
- HW Virtualization support (AMD “Pacifica”)

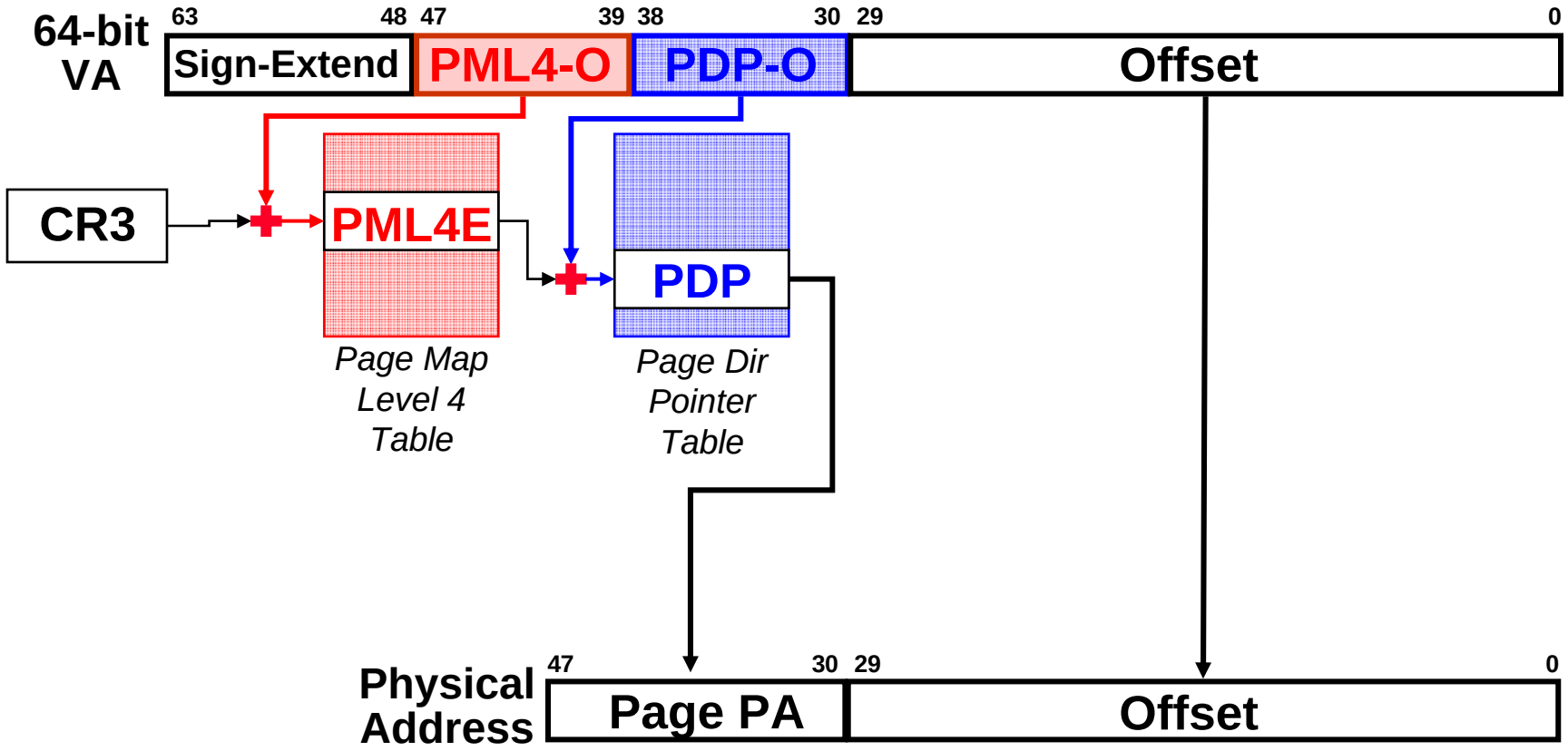
RDTSCP: Read Serialized TSC Pair

- New instruction, similar to RDTSC:
  - Returns 64-bit TSC value in %edx:%eax
  - Is a serializing operation -- prevents speculative reads of TSC
  - Returns TSC_AUX[31:0] MSR in %ecx at same time as TSC
    - *OS initializes TSC_AUX to meaningful value*
    - *Atomicity ensures no context switch btw read of TSC & TSC_AUX.*
  - Availability determined by new extended CPUID feature flag
- Allows TSC and OS-supplied value (such as CPU number) to be read atomically in a serializing way in user mode.
  - TSC rates between CPUs in MP-system may vary
  - Linux can put CPU number in TSC_AUX so user-mode get-time-of-day knows which per-cpu adjustments to use to fix-up TSC value.

Planned 2007 Processor Features

- Multi-core capable
- DDR3 support
- 1-GB pages – see next slide
- 48-bit Physical Addressing – see later slide
- Greater than 32-socket support
- P-state Invariant TSC (APIC Timer is already)
- P-state Fire-n-Forget
- Monitor/Mwait
- Shared L3-cache
- Further Virtualization extensions

1 Gigabyte Pages & 48-bit Physical Addresses



Plan is for Physical Address in PTEs to be 48 bits for all page sizes.

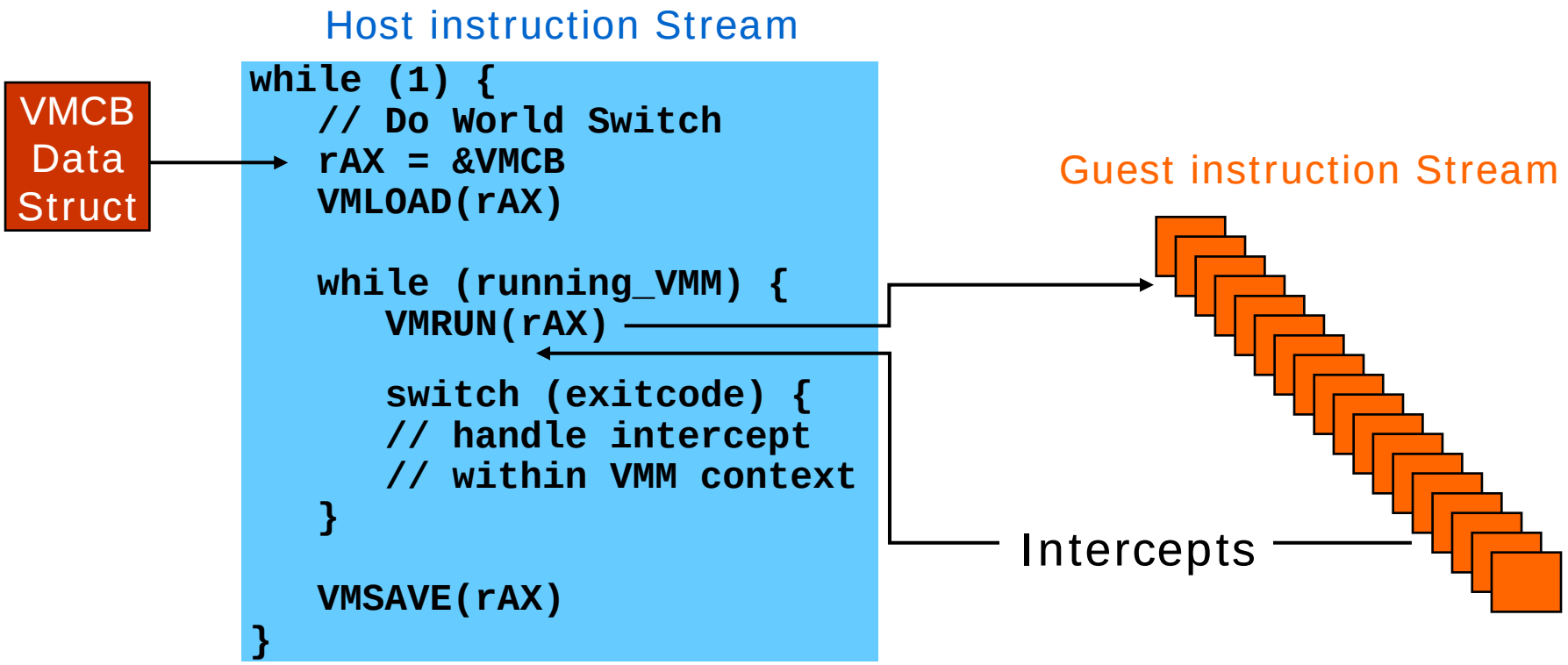
Virtualization Discussion

AMD Virtualization Directions

- AMD “Pacifica”: HW-Virtualization-Assist. Base features planned launch in 2006 Generation
- Primary components of Architecture:
  - Host/guest management hardware support
  - Event Injection
    - *Eliminates need for VMM code to emulate x86 exception delivery*
    - *Designed to reduce VMM development time significantly*
  - Nested Page Tables
    - *Designed to improve VMM performance, and reduce overhead*
    - *Helps reduce VMM complexity*

Core “Pacifica” Architecture: VMRUN

- Virtualization based on Virtual Machine Run (VMRUN) instruction
- VMRUN executed by host causes the guest to run
- Guest runs until it exits back to the host
- Host resumes at the instruction following VMRUN

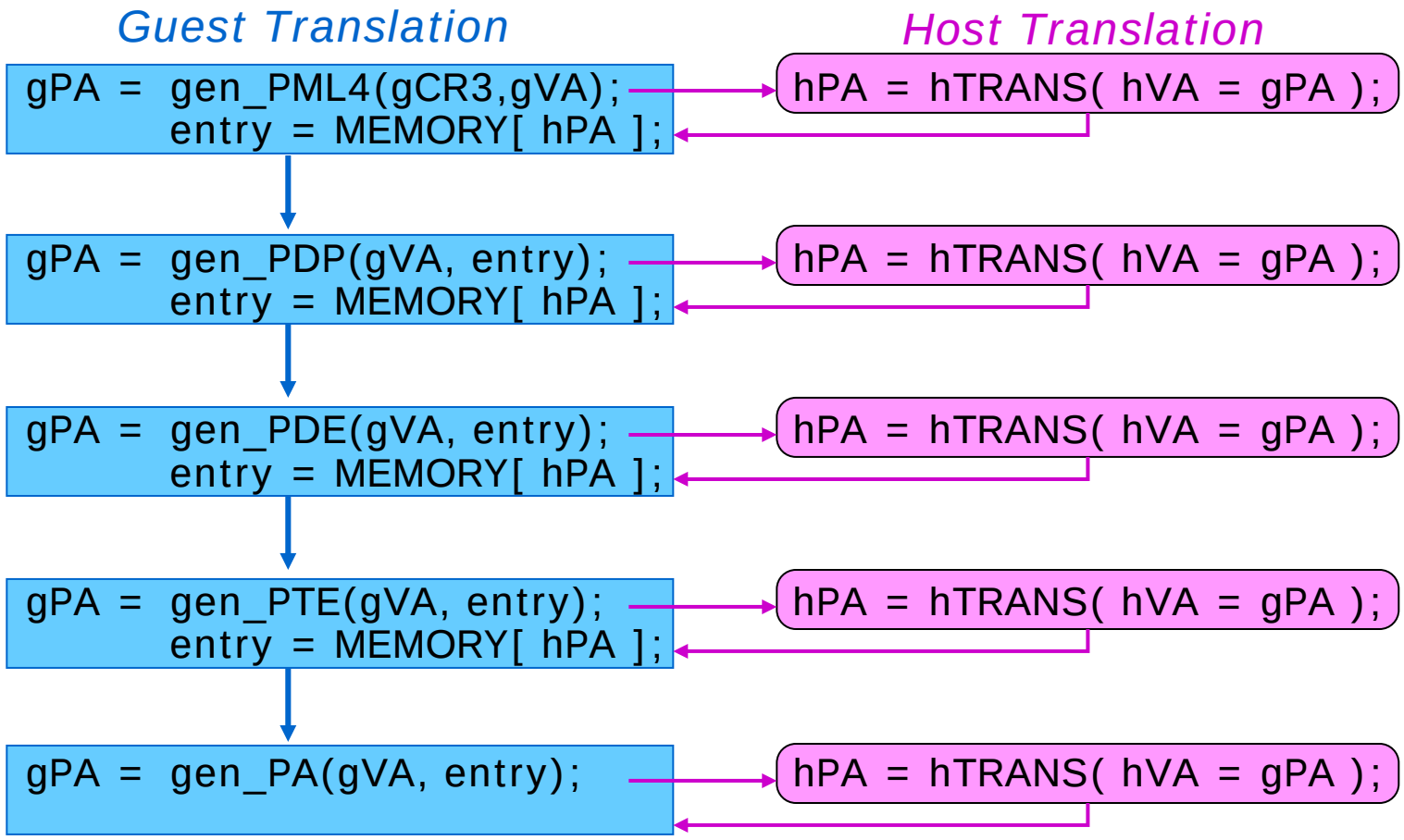


Core “Pacifica” Architecture: Intercepts

- Guest runs until:
  - It performs an action that causes an exit to the host
  - It explicitly executes the VMMCALL instruction
- The VMCB for a guest has settings that determine what actions cause the guest to exit to host
  - These intercepts can vary from guest to guest
  - Two kinds of intercepts
    - *Exception & Interrupt Intercepts*
    - *Instruction Intercepts*
  - Rich set of intercepts allow the host to set customize each guest’s privileges
- Information about the intercepted event is put into the VMCB on exit

Nested Paging

- CPU maps each Guest_PA to Host_VA and then translates to Host_PA
- CPU builds compound gVA_to_hPA TLB entries (guarded by ASID)
- Far more efficient than “Shadow Page Tables”, all handled by CPU

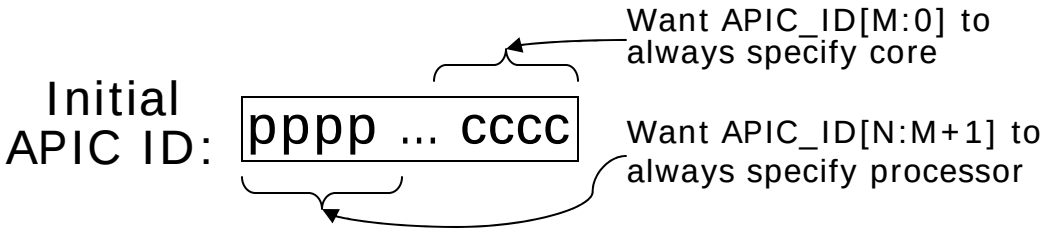


Challenges / Issues

Multi-core Numbering

- Assume system has non-power-of-two number-of-cores in at least 1 processor due to design or retirement of bad core(s).
  - How to tell OS? How to keep "sanity" in core/processor bit masks?
- BIOS calculates "Rounded Number of Cores" (RNC):
  - *RNC = 2^ceil(log2(Number_of_Cores))*
- BIOS assigns APIC IDs of each processor's cores to an RNC-aligned block of IDs:
  - *APIC_ID[proc=i, core=j] = RNC * (OFFSET + i) + j*

- Example: 2-processor system
  - proc 0 has 3-cores
  - proc 1 has 4-cores
  - RNC = 4 on all cores



Proc	Core	APIC ID	
0	0	0x8	= 4*(2+0) + 0
0	1	0x9	= 4*(2+0) + 1
0	2	0xA	= 4*(2+0) + 2
		rsvd	*= 4*(2+0) + 0*
1	0	0xC	= 4*(2+1) + 0
1	1	0xD	= 4*(2+1) + 1
1	2	0xE	= 4*(2+1) + 2
1	3	0xF	= 4*(2+1) + 3

Multi-core Numbering (cont)

- OS should use same process to discover topology of processors & cores.
- OS can not assume that BSP's CPUID.number_of_cores is same for all processors.
- OS can assume that RNC calculated on any processor is same for all processors.
- BTW, APIC ID and IOAPIC ID space are disjoint spaces, so the IDs can overlap
  - There is no logic driven by IOAPIC ID on AMD K8 chipsets

Other Challenges

- Mixture of physical and logical cores
- UEFI and other BIOS Adventures
- PCI Express® Extended Configuration Space
- Std Kernel-supported Performance Counter mechanism
  - don't steal NMI
- 1-GB Page support
- Number of threads per-process
- HW-assisted Virtualization

Trademark attribution and cautionary statement

CAUTIONARY STATEMENT

This presentation contains forward-looking statements, which are made pursuant to the safe harbor provisions of the Private Securities Litigation Reform Act of 1995. Forward-looking statements in this presentation relates to, among other things, product and technology introduction schedules; and planned product architecture features. Forward-looking statements involve risks and uncertainties that could cause actual results to differ materially from the company's current expectations. Risks include the possibility that the company may not achieve its current product and technology introduction schedules; that global business and economic conditions will worsen resulting in lower than currently expected sales; that Intel Corporation's pricing, marketing programs, product bundling, new product introductions or other activities targeting the company's processor business will prevent attainment of the company's current plans; that demand for personal computers and, in turn, demand for the company's processors, will be lower than currently expected; that adoption of AMD64 products by OEMs will not continue as expected; that the company will not be able to meet demand for its products; that the company will not be able to raise sufficient capital to enable it to establish leading-edge capacity to maintain its market leadership positions; and that solutions providers will not timely provide the infrastructure to support the company's AMD64 technology.

Because the company's actual results may differ materially from our plans and expectations today, we encourage you to review the company's filings with the Securities and Exchange Commission, including but not limited to the risks and uncertainties contained in our Annual Report on Form 10-K for the year ended December 26, 2004, and our Quarterly Report on Form 10-Q for the quarter ended March 27, 2005.